

(21) Application No **0115461.6**

(22) Date of Filing **22.06.2001**

(71) Applicant(s)  
**Hewlett-Packard Company**  
(Incorporated in USA - Delaware)  
3000 Hanover Street, Palo Alto,  
California 94304, United States of America

(72) Inventor(s)  
**Colin F Anson**

(74) Agent and/or Address for Service  
**Robert F Squibbs**  
**Hewlett-Packard Limited,**  
**Intellectual Property Section, Filton Road,**  
**Stoke Gifford, BRISTOL, BS34 8QZ,**  
**United Kingdom**

(51) INT CL<sup>7</sup>  
**G06F 17/30**

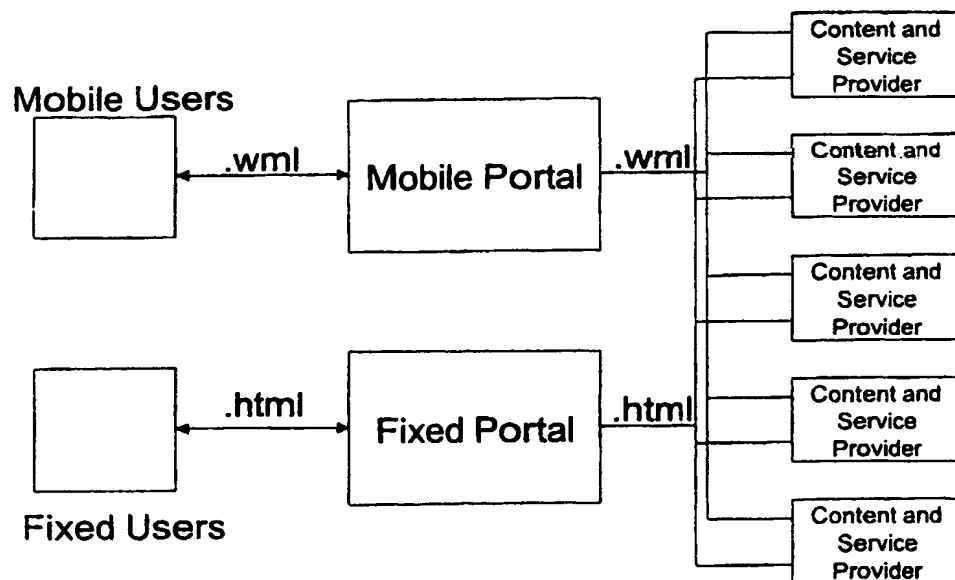
(52) UK CL (Edition T )  
**G4A AUDB**

(56) Documents Cited  
**GB 2346238 A** **WO 2001/077842 A1**  
**WO 2001/076190 A2** **WO 2001/065355 A1**  
**WO 2001/042989 A2** **WO 2001/035235 A1**

(58) Field of Search  
INT CL<sup>7</sup> **G06F 17/21 17/22 17/30**  
Other: **ONLINE: EPODOC, JAPIO, WPI**

(54) Abstract Title  
**Portal to allow access to web pages using different formats**

(57) A content-aggregating portal is provided for aggregating content from multiple content providers that each make available content in XML format. The portal as well as aggregating content as appropriate to satisfy user requests, also uses XML style sheets to apply transformations to the XML-based content to effect adaptations such as user device adaptation (Eg WAP, PDA, DESKTOP), user preferences adaptation, or location-based adaptation. The content providers provide their content according to a common XML semantic so as to enable the portal to provide consistent user responses when aggregating from several content providers. The common XML semantic may be specified by the portal operator or informed by related actions as specified in a services registry. In a simpler embodiment, a content-indexing portal is provided.



**Figure 1**

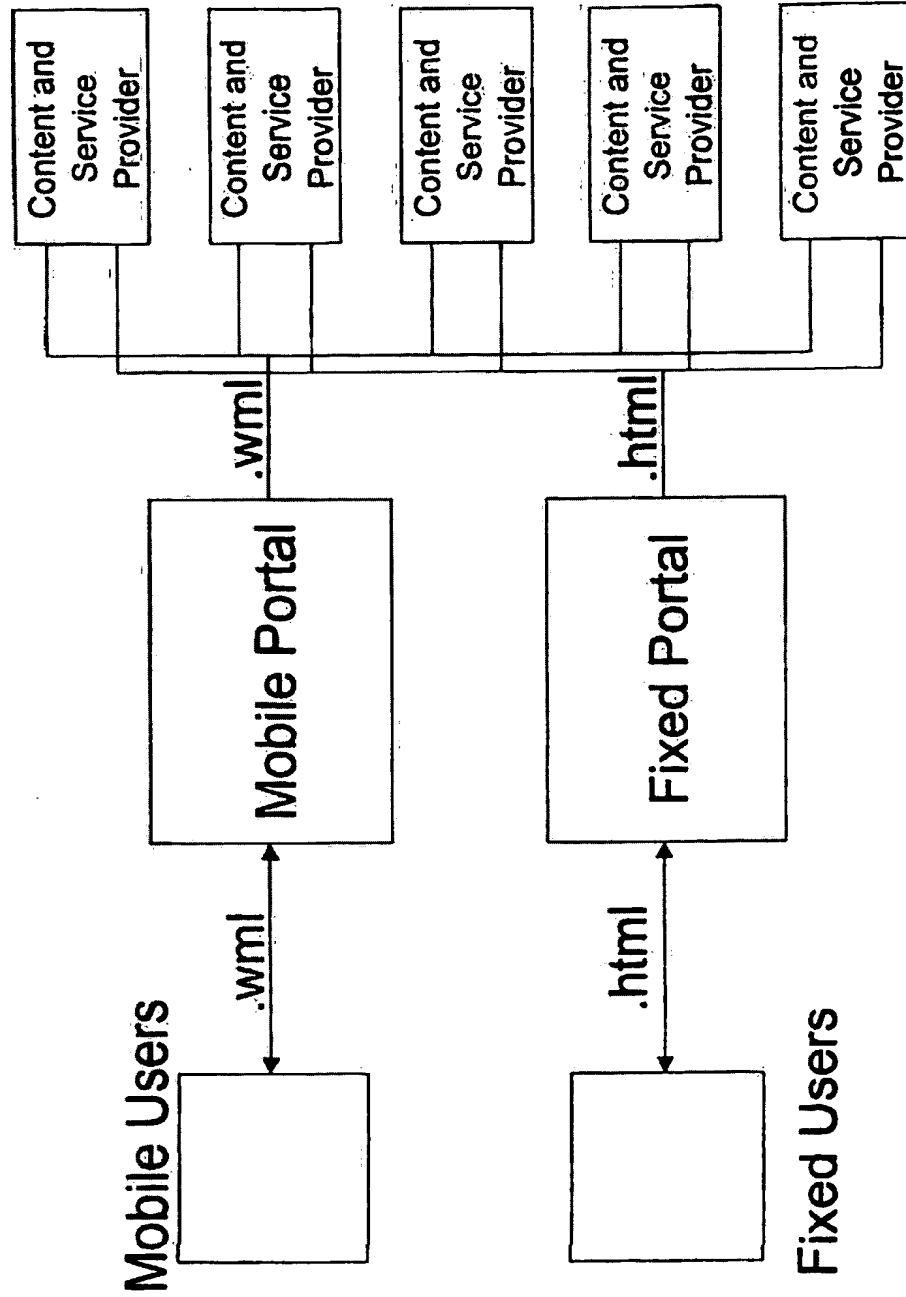


Figure 1

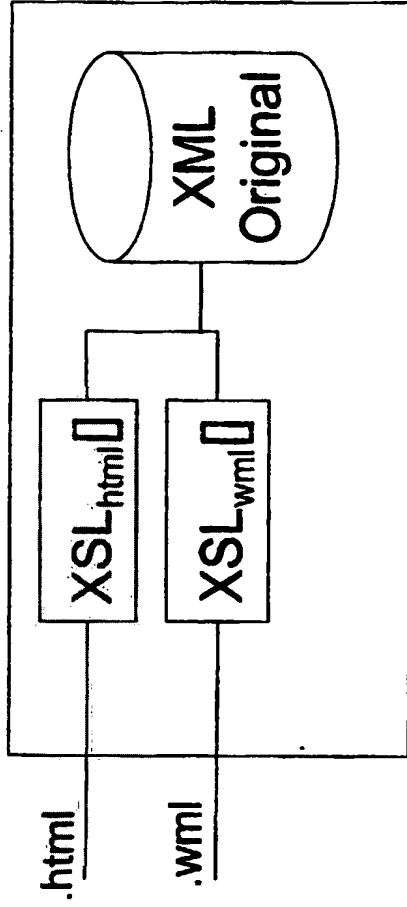


Figure 2

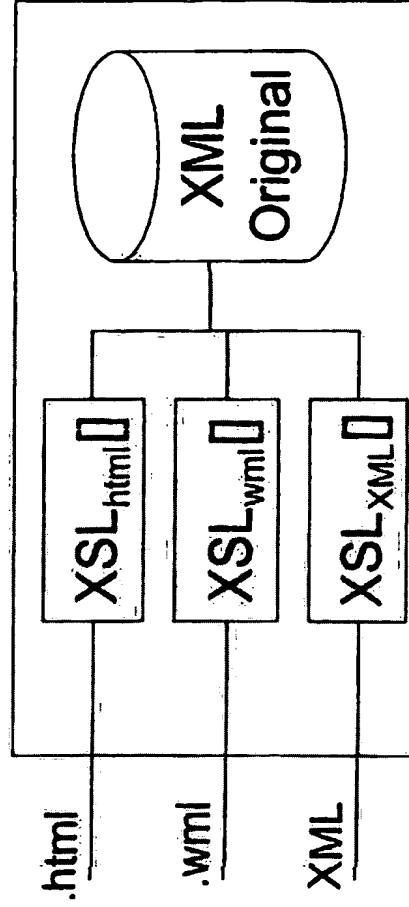


Figure 3

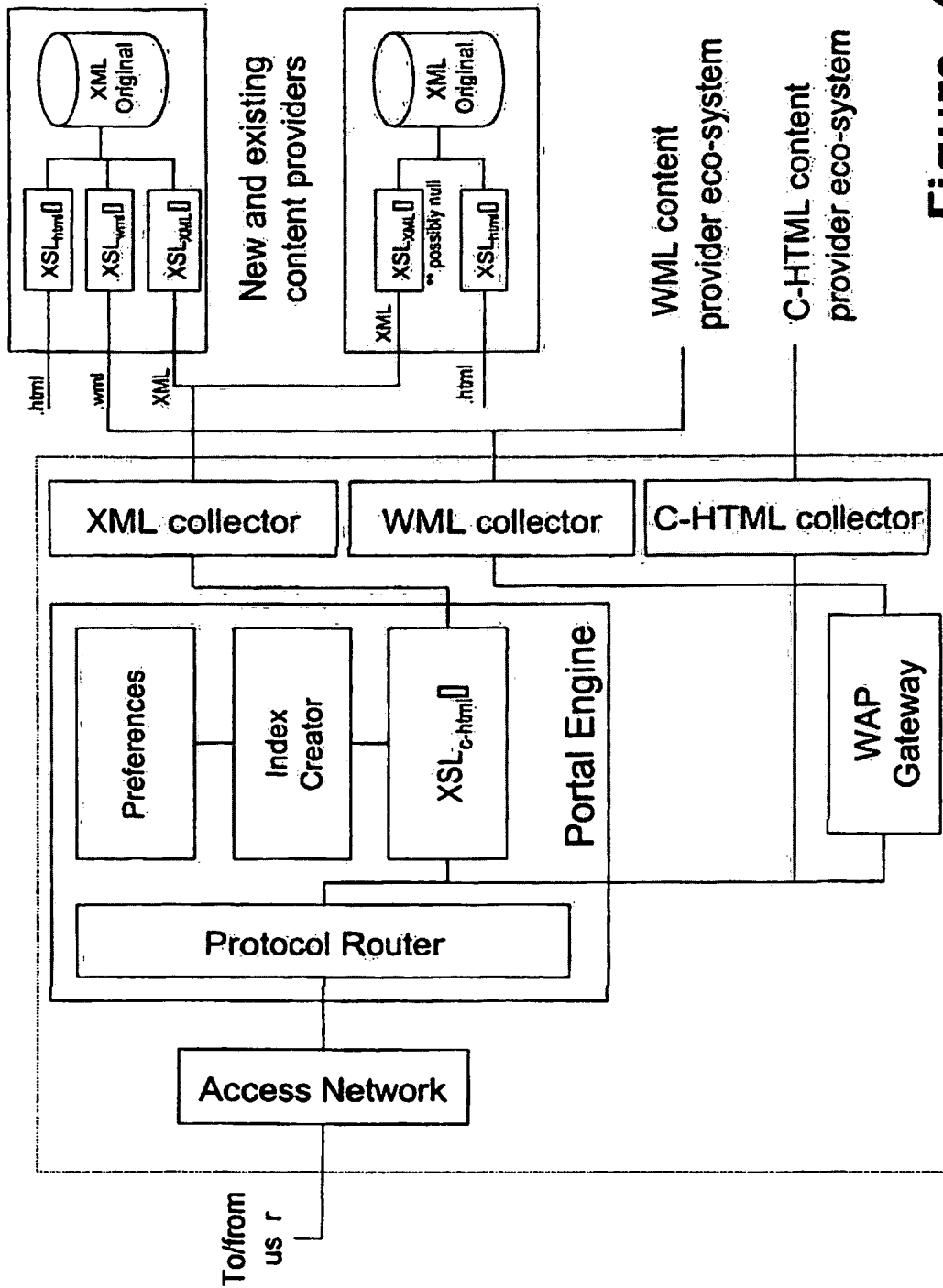


Figure 4

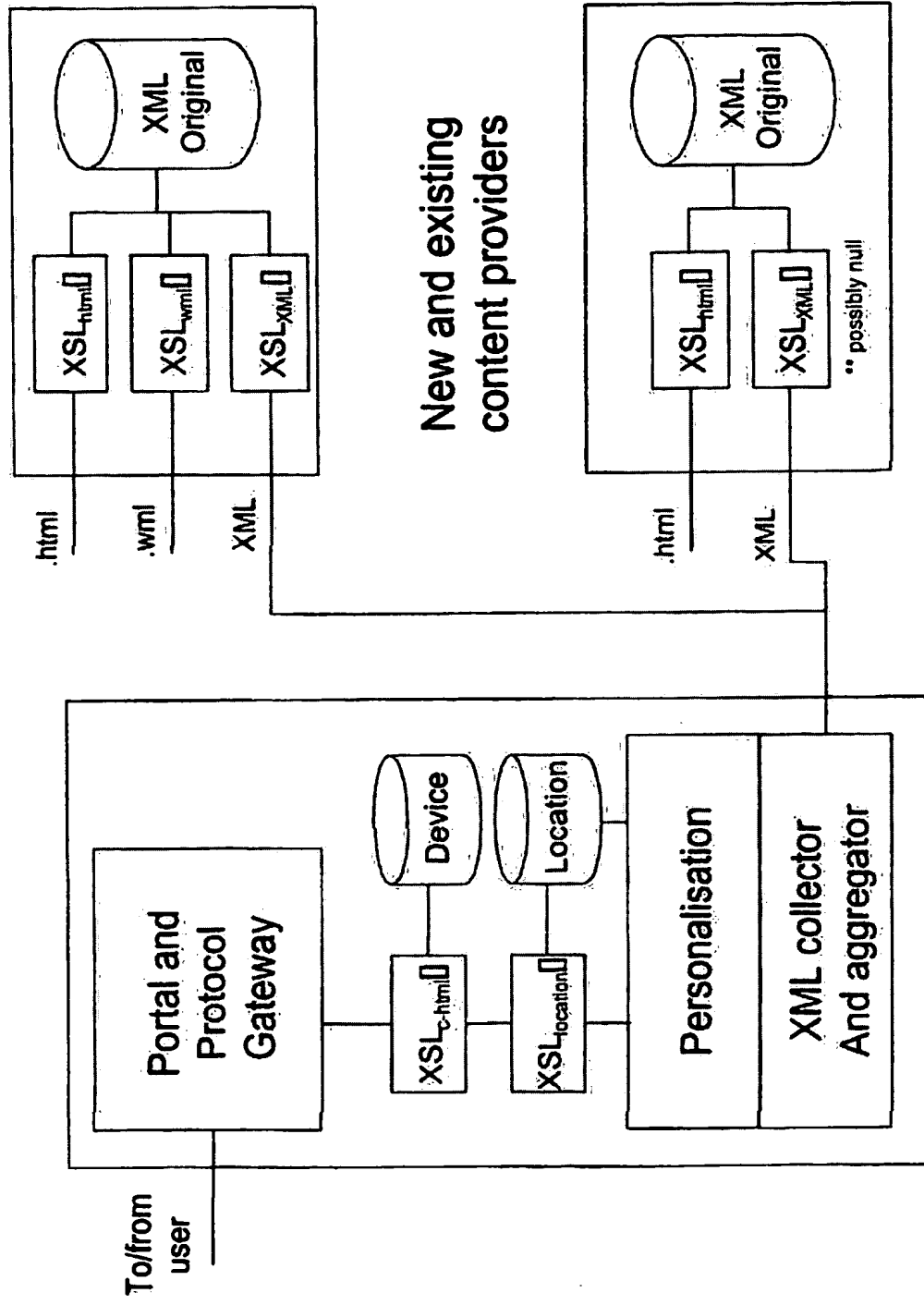


Figure 5

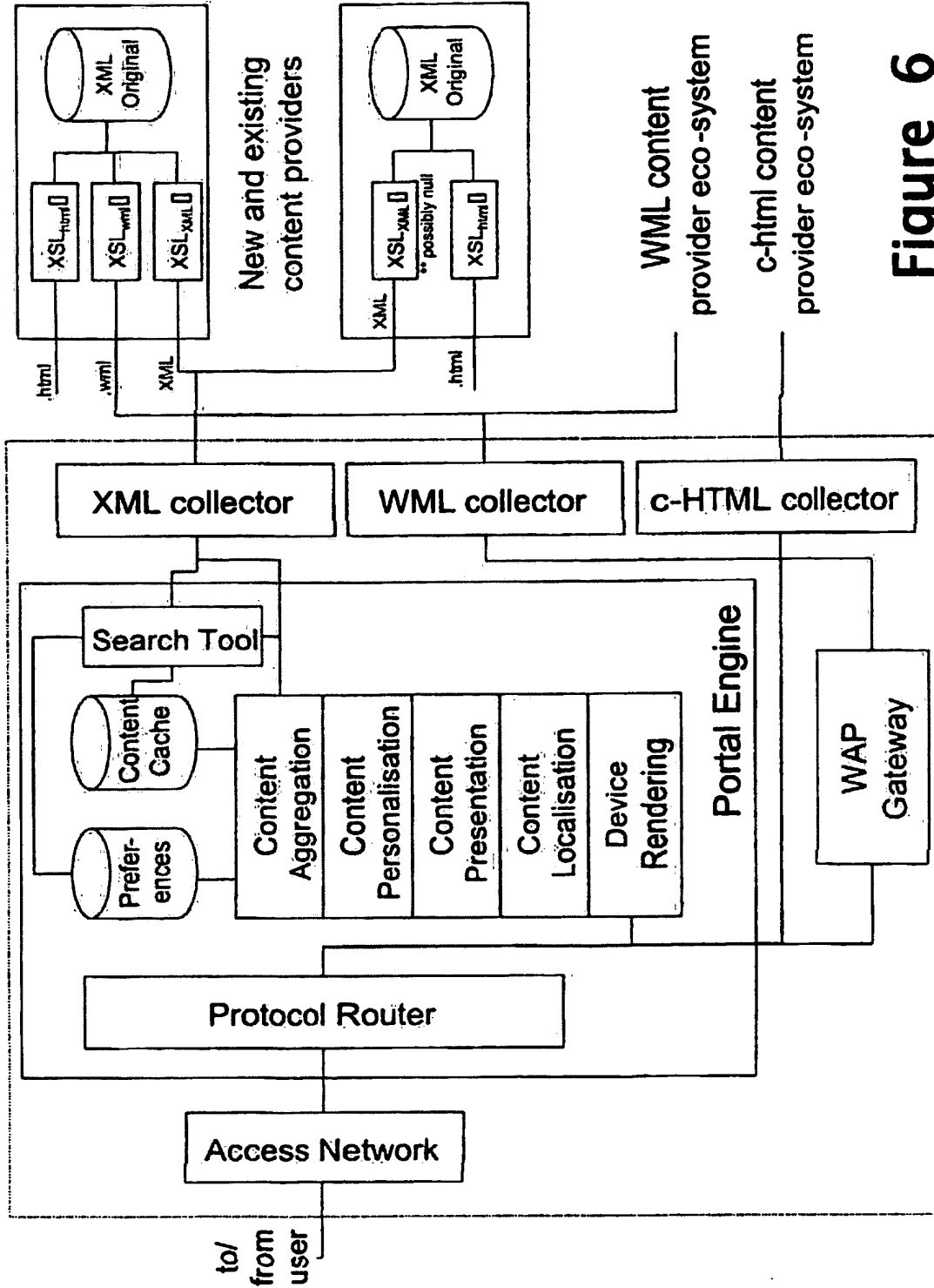


Figure 6

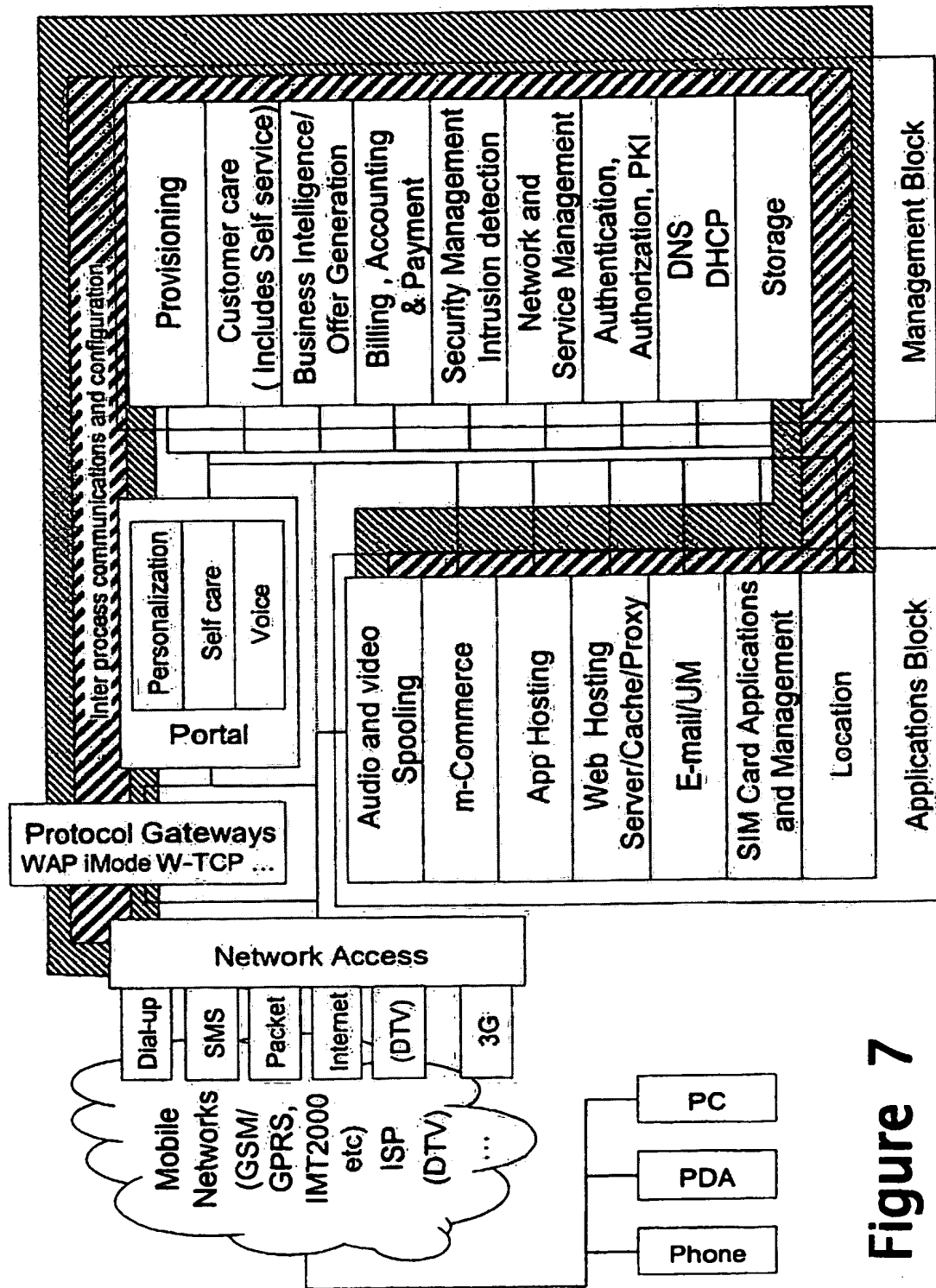
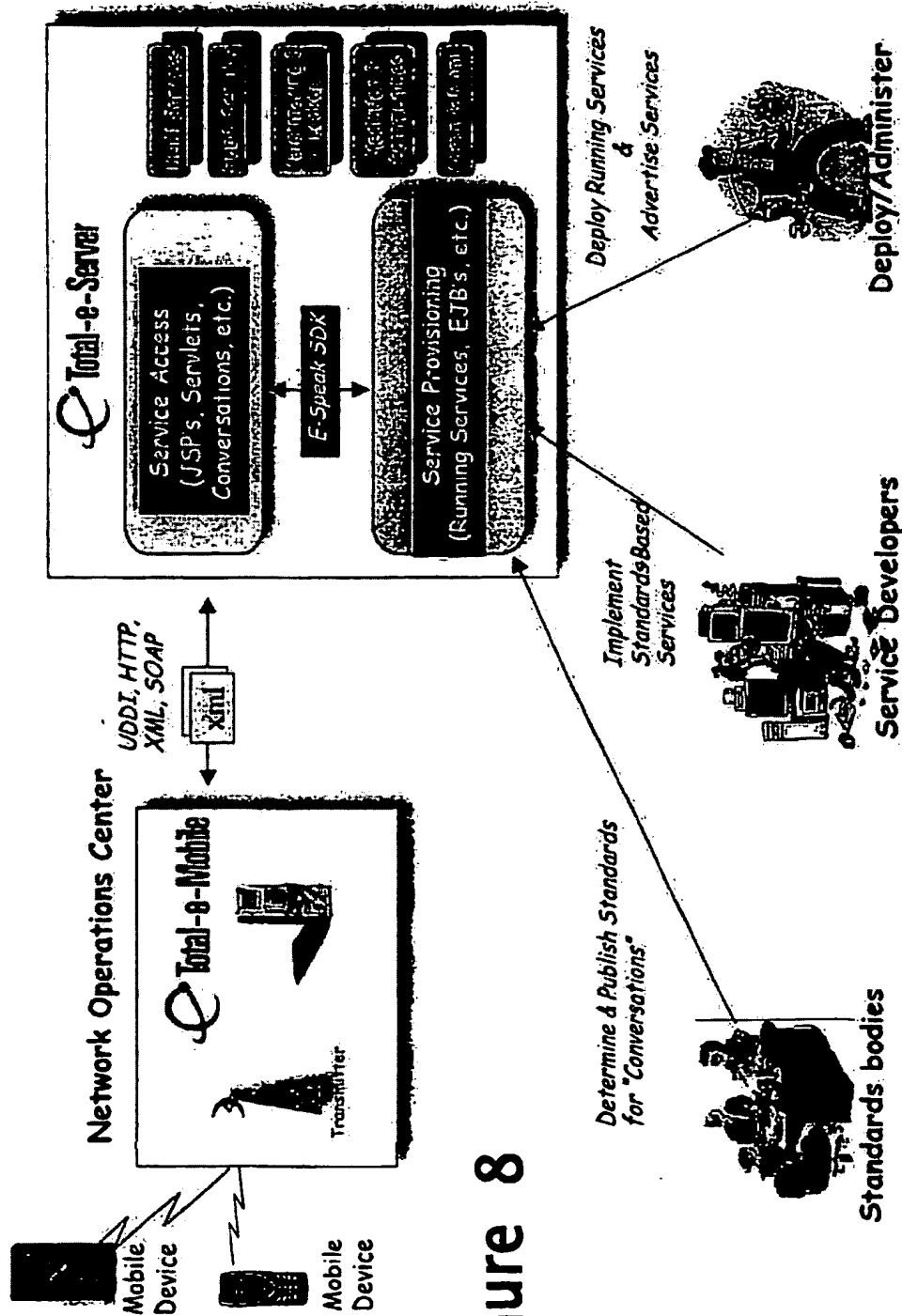


Figure 7



El Camino Architecture, Author: MobileTEC team

Figure 8



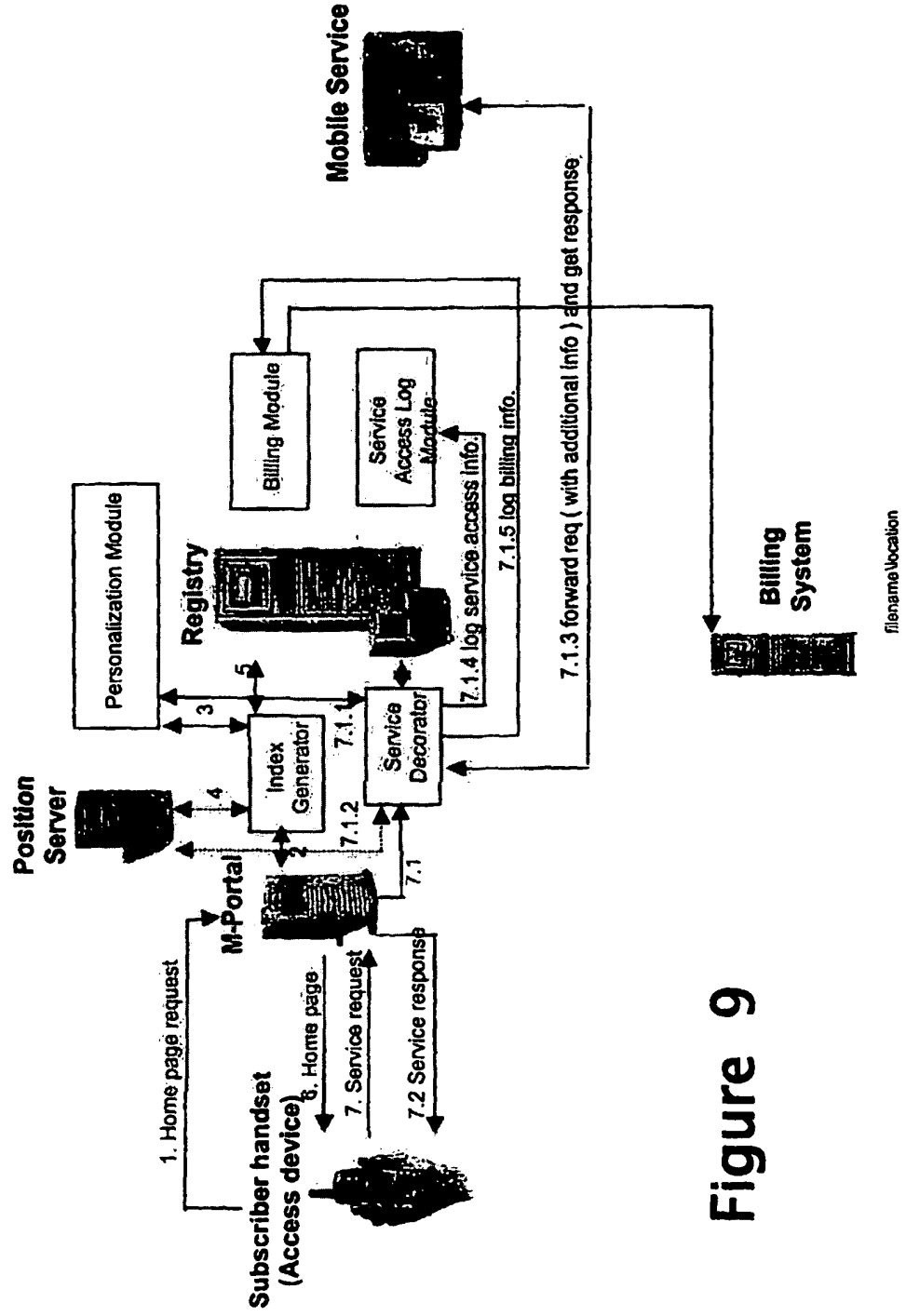


Figure 9

## **Distributed Content Indexing and Content Aggregating Portals**

### **Field of the Invention**

The present invention relates to distributed content indexing and content aggregating  
5 portals

### **Background of the Invention**

Today, many mobile services have to be implemented by taking information initially  
created for presentation focused .html documents and converting to .wml ("Wireless  
10 Markup Language", part of the WAP - Wireless Application Protocol - specification). This  
approach results in the content providers making available two different forms of each  
document, one for fixed users (html documents) and the other for mobile users (wml  
documents) - see Figure 1 of the accompanying drawings. While utilities exist to automate  
this conversion it is not possible to fully automate a guaranteed successful solution. Thus,  
15 for example, Everypath has a clever solution where humans are used to identify what parts  
of web pages need to be conveyed to a mobile user. These are marked and then the system  
performs the conversion in real-time. This can work well until the original web site is re-  
designed. Other, more automated solutions, typically disappoint. The result is that good  
presentation in .wml comes from good original .wml encoding, done specifically for  
20 mobile and therefore a dual (fixed and mobile) content eco-system exists. Here, the  
original information provider is adversely impacted by the diversity.

In Japan, NTT DoCoMo is currently running a very successful mobile e-services business  
in Japan based on "i-mode" which can be characterized as a combination of technological  
25 and business strategies that together give mobile users a very good information-access  
experience. I-mode uses a simplified form of html for content known as "c-html". One of  
the important business components has been to assemble a very large amount of valuable  
content, adapted for presentation in a tightly specified environment, to form the core of an  
offering. The original, critical mass, core pulled in thousands of other content providers to  
30 create a single very viable eco system. While c-html could be deployed in other markets it  
would not have the same impact simply because the challenge to build a critical mass  
would be huge. As already indicated, these markets already suffer from multiple content

types with web, WAP and sometimes SMS forms. Content providers find this situation unacceptable and will resist all suggestions of another presentation focused content encoding, especially with the eventual practical requirement to migrate to XHTML. XHTML, sometimes called the future of the Internet, obeys the grammar rules of XML and conforms to the vocabulary of HTML. Presentation focused, it is expected that XHTML will be widely adopted for content that is presented on PC browser applications. As such, automated content conversions from material destined for big screens to small screens will be easier. But, it does not address the fundamental concern over how to automate the impact of content presentation equally on a large PC screen as well as a small one on a mobile phone. This is recognised in the creation of basic XHTML with a restricted set of capabilities for small screen devices

In fact, XML can be used as a potential single source of .html and .wml documents. Content providers can automatically create the various encodings from a single source using XML style sheets. These XML style sheets can be used to select items from a document and convert the document to another form. For example, they can be used to put device specific formatting in the document as well as filter a document to provide only location specific information.

The logic of style sheets can be represented as follows:

$$\text{XSL}[\text{XML}] \rightarrow \text{XML}$$

that is, a style sheet used on an XML document can produce another XML document

This notation can be used to show conversion style sheets;

$$\text{XSL}_{\text{html}}[\text{XML}] \rightarrow \text{.html}$$

that is, a style sheet used on an XML document can produce an html document

Similarly:

$$\text{XSL}_{\text{wml}}[\text{XML}] \rightarrow \text{.wml}$$

that is, a style sheet used on an XML document can produce a wml document

Hence to perform localisation and device type characterization, the following operation

can be performed:

$$\text{XSL}_{\text{wml}}[\text{XSL}_{\text{device\_type}}[\text{XSL}_{\text{position}}[\text{XML}]]] \rightarrow \text{.wml}$$

- For a content provider who has the sources of information in XML, and who wants to produce good Web and WAP services, this is a good starting point. One encoding and then multiple style sheets produces all the content type encodings. They could even add a c-html style sheet. While there are some shortcomings with a pure XML story, good control of all aspects of information creation as well as style sheet creation will get round them. Specifically, there are problems in the area of semantic meaning beyond that possible with locally defined DTDs (Document Type Definitions). Furthermore the content provider takes responsibility to directly provide the content, notably the presentation, to the mobile and therefore the role of a portal/intermediary, like an i-mode service, is simply that of an index.
- A content provider with just html has a number of migration routes to provide mobile content. The content provider could provide parallel .html and .wml as described above. Alternatively the content provider could go directly to a core XML original and use stylesheets to create the .html and .wml (see Figure 2)
- It is an object of the present invention to provide improved ways of providing content to users through portals.

### **Summary of the Invention**

- According to one aspect of the present invention, there is provided a method of providing content to users via a communication infrastructure, wherein a plurality of content providers each make available content in XML format, and an portal operator provides a content-indexing portal through which users can access content from the content providers, the portal using XML style sheets to apply at least one of the transformations to XML-based content retrieved from a content provider in response to a user request:
- user device adaptation of the content;
  - user preferences adaptation;
  - location-based adaptation;

the content providers providing their content according to an agreed semantic meaning for the XML whereby to enable the portal to provide consistent user responses in respect of content from said plurality of content providers.

5 According to another aspect of the present invention, there is provided a method of providing content to users via a communication infrastructure, wherein a plurality of content providers each make available content in XML format, and an portal operator provides a content-aggregating portal through which users can access content from the content providers with the portal serving to aggregate content from multiple content  
10 providers as appropriate to provide a response to a user request, the portal using XML style sheets to apply at least one of the transformations to XML-based content retrieved from a content provider in response to a user request:

- user device adaptation of the content;
- user preferences adaptation;
- 15 - location-based adaptation;

the content providers providing their content according to an agreed semantic meaning for the XML whereby to enable the portal to provide consistent user responses when aggregating from multiple said content providers.

20 According to a further aspect of the present invention, there is provided a method of providing services to users via a communication infrastructure, wherein a plurality of service providers each make available content in XML format associated with services, and an portal operator provides a portal through which users can access content from the service providers, the portal using XML style sheets to apply at least one of the  
25 transformations to XML-based content retrieved from a content provider in response to a user request:

- user device adaptation of the content;
- user preferences adaptation;
- location-based adaptation;
- 30 the service providers providing their content according to an agreed semantic meaning for the XML whereby to enable the portal to provide consistent user responses in respect of services from said plurality of service providers.

### **Brief Description of the Drawings**

- 5 Embodiments of the invention will now be described, by way of non-limiting example, with reference to the accompanying diagrammatic drawings, in which:
- . **Figure 1** is a diagram of a prior art approach to content production for web and mobile users;
  - . **Figure 2** is a diagram an alternative prior art approach to content production for web  
10 and mobile users;
  - . **Figure 3** is a diagram illustrating the provision of content in a standard form XML;
  - . **Figure 4** is a diagram of a content-indexing portal embodying the invention;
  - . **Figure 5** is a diagram showing XML data flows for a content-aggregating portal embodying the invention;
  - 15 . **Figure 6** is a functional block diagram of the portal subject of Figure 5;
  - . **Figure 7** is a diagram a the Hewlett-Packard Mobile E-Services Architecture;
  - . **Figure 8** is a diagram of a services portal embodying the invention;
  - . **Figure 9** is a diagram illustrating a service access operation of the Figure 8 services portal.

20

### **Best Mode of Carrying Out the Invention**

#### **Establishing Content Eco-System Control**

- 25 To provide a coherent content eco-system, existing XML users and html core users, would be encouraged to make their content available as XML as depicted in Figure 3. The operator (or possibly another party) would ensure the meaning conveyed in the XML is consistent and would define a single standard for the vocabulary, for example be providing or specifying a small set of Document Type Definitions (DTD). For simple services, at  
30 start-up, this could be a few pages in length. Existing XML users would require an additional style sheet to convert the XML documents, specifically to assign the right tag names. Core html users would need to be encouraged to export their content in XML.

While they could invent their own XML structures, the operator standard would provide a very attractive starting point. Few other companies could take this role.

The purpose of XML in this arrangement is to give a structured document that can easily be searched and later on accessed for content aggregation. There are advantages to the operator taking the responsibility for the meaning in the vocabulary and this could, at least in part, be done by referencing existing and upcoming XML based standards that cover difference aspects of content syndication, including NewsML and NITF ([www.iptc.org](http://www.iptc.org)) and PRISM ([www.prismstandard.org](http://www.prismstandard.org)). Similarly there are standards dealing with the transfer and management of syndicated content like the Information and Content Exchange protocol ([www.icestandard.org](http://www.icestandard.org)).

The benefit of the proposed arrangement is that search engines can deliver more consistent and hopefully accurate results, improving the user experience.

15

For delivery to a mobile, this proposed arrangement is not technically constrained to the specifics of content encoding for i-mode or WAP. Equally, the following conversion are possible:

`XSLwml[XSLdevice_type[XSLposition[XML]]] -> .wml`

20 and

`XSLc-html[XSLdevice_type[XSLposition[XML]]] -> .html (of type c-html)`

This gives the foundation for a solution where the content providers export XML and the mobile operator's servers perform the final style-sheet manipulations, such as device type manipulations. This is advantageous as only the operator need hold the entire set of device specific style sheets, not the content provider. This creates a simple but effective control point, through the ownership of 'meaning' and device formatting, to impose business rules to ensure a good user experience. Furthermore, by making the conversion style sheets available to content providers, they can check how their content will be converted and presented.

30

### Content Indexing Portals

The foregoing approach enables the provision of content-indexing portals giving improved end-user experience.

- 5 A foundation for a simple content indexing portal lies in the business rules for content providers and the content translations executed in the portal. Figure 4 illustrates an embodiment of a content-indexing portal working on these principles. The portal basic service is a personalised index for each mobile user that adapts with usage and specifically points to content the user has indicated they are interested in.

10

The portal also controls content presentation. A clear knowledge of the XML stylesheet rules permits an XML content creator to be sure on how their content will appear yet they do not need to create or buy all the potential XML translation style sheets.

- 15 With XML based content and the final mobile content encoding in the portal, migration to WAP-NG is integrated in the design.

To ensure a good user experience, the content indexing portal relies on rules that state:

- The limited set of screen sizes and colours that are supported - device profiles
- 20 • How content is obtained from content providers (protocol) - content providers will be required to export content in the agreed XML formats using http or https, where the XML documents appear simply as another mime type.
- How the content providers provide the content in a consistent manner - there will be one standard meaning for the XML vocabulary for all content; to achieve this operator
- 25 (or possibly another party) will define, in simple terms, the meaning for all content types.

- The portal/server, owned by the network operator, will then provide basic index page services exactly as in i-mode or simple WAP portals and the content will be collected in
- 30 XML. This will be transformed in the portal/server to create device specific content using XML style sheets (and other tools).



The portal/server may reference and forward content that is already formatted for WAP or c-html presentation. This will be transparently passed through the portal as required.

Based on XML, with the agreed operator vocabulary meanings, the content indexing portal  
 5 personalises and renders personal service menus. Localisation could easily be added. These web assets are then delivered by references to the services themselves.

#### Content and Service Aggregating Portals

10 A content-indexing portal is in many ways only a partial solution to user needs – what is often really called for is an information and service aggregation portal where information is pulled together for a user rather than simply referenced.

Figure 5 depicts, in simplified form, a content-aggregating portal in terms of the XML  
 15 document flows for a c-html user-device environment. XML documents, in the agreed vocabulary, are collected by the portal. The content is then manipulated in various ways - aggregated, personalised, localised and converted. The diagram shows XSL stylesheets used for location and device type conversions. (The diagram only shows the XML paths, it does not show the paths required to relay original WAP and i-mode information without  
 20 any manipulation.). Figure 6 is a functional block diagram of the content-aggregating portal.

The detailed implementation of the portal will depend on requirements for performance. One issue is whether the content is always to be collected fresh from the content provider  
 25 when required or whether the content to be cached and the cache refreshed from time to time. Equally, it must be decided whether aggregation is to be done on demand or in advance. In a preferred embodiment, a 2 level content cache is provided, holding both raw content and pre-aggregated content.

30 Whilst the content-aggregating portal could retrieve content based on pre-established indexes, it is preferred to provide a search engine to ensure that the most appropriate content is obtained. This search engine works in 2 ways

- specifically at the request of the user  
speculatively in searching for the best content for typical requests and caching the content

A filtering mode can also be implemented where new content is continuously evaluated  
5 against individual user profiles. The filtering mode can, for example, be supported by the  
Chilead “Current Awareness” solution and allows users to be proactively notified about  
new interesting content via, for example, SMS. In Figure 5, this could happen inside the  
“XML collector and aggregator” block.

10 In a content aggregating portal, the personalisation, location and content aggregation  
systems directly interact. With a consistent representation of content, management and  
aggregation is possible using standard ISV products including Interwoven, Broadvision and  
ATG Dynamo.

15 The Figure 5 content-aggregating portal provides the following benefits:

- It integrates together valuable content in a manner to ensure a good user experience.
- It does not impose technical restrictions on hybrid WAP and i-mode solutions.
- It contains the foundations for extension into deeper content aggregation

20

The content aggregating portal can be deployed using the Hewlett-Packard Mobile E-  
Services Architecture and specifically, the HP Bluestone J2EE and XML environment as  
well as Total-e-mobile. The HP Mobile E-Service Architecture is depicted pictorially in  
Figure 7.

25

#### A Generic Approach with Migration into E-Services

The above-described content aggregating portal is a distributed XML document integration  
bus that uses standard web protocols but which had the operator effectively assuming the  
role of a standards body; while this is satisfactory in a single national market, it is less  
30 likely to work in a multinational environment.

Equally, the focus of the above has been on presenting content rather than delivering useful

services. In a mobile world, there will be a greater need for an activity, that is, a need to actively respond to "Change the time of my flight" by doing just that rather than showing a list of alternative flight times; also, changing the flight time is likely to have knock-on effects requiring action to change car hire booking, late arrival at the hotel, etc.

5

To resolve these issues a more generic portal architecture is envisioned. In this case rather than define a single set of meanings for the XML based content, a service registry is set up that contains many service offers, each using a published vocabulary with well understood meaning and purpose. The solution is now universally deployable without requiring the operator to become a standards authority rather, to become the holder of the critical services repository. Furthermore the solution is based on delivering e-services, i.e. executing activities, rather than reading web pages.

10

To achieve this, each service offer must be wrapped in a well-defined manner. Each e-service performs a task and can be registered, discovered and used. This is the underlying logic of HP's e"speak architecture. Indeed, these e-services can be contents, activities, operator-services in the portal, like billing and location, or any other asset that can be placed on the Web.

15

The generic form of the architecture could utilise HP e"speak, creating a mobile e-service integration framework (see Figure 8). This can be deployed with the HP Bluestone Total-e-Mobile to create a whole solution, which can be hosted in the HP Mobile e-Services Architecture.

20

A service integration framework can be based on HP e"speak, where all Web based assets are treated as e-services. Each is described in the registry and a composite e-service is created from the individual services. The portal, with service aggregation, would be constructed as a set of services, each delivering exactly the same logical blocks as before but rather than simply pulling XML documents from remote sites, the remote sites would offer their assets as an e-service. They too would wrap their assets in the well-defined manner, specified via HP e"speak, and their content appears as a service - a description of the weather that can be animated rather than textual description.

30

The simple example shown in Figure 9 shows a step-by-step operation of the service. It assumes that each service has already been registered in the Registry.

- 5 In this example, a user requests from the portal their home page (1). The portal calls on the index generator (2), which collects personalisation (3) and location (4) information before going to the Registry to collect a list of services (5) for the home page (6). When a service is requested (7) a service decorator is used to pull together the request (7.1.1, 7.1.2), which is forwarded with additional useful parameters (7.1.3.). Once the service responds then
- 10 event logging (7.1.4) and billing (7.1.5) can be carried out before the final delivery (7.2).

The example shows different types of resource that are wrapped as e-services and illustrates re-use of the personalisation and location servers. These are operator provided services that are used in conjunction with third party services to deliver the whole user

- 15 experience. Clearly a new service can be rapidly created by re-use of the e-services that are known by the Registry.

In the foregoing, a three-phase architecture has been described progressing from content indexing portals, through content aggregating portals into portals that support activity

- 20 based e-services. The content semantics problem can be resolve in 2 ways - eco-system control over the meaning of XML content or by use of a Registry where the meaning of the content is clarified by the action. It will be appreciated that many variants are possible to the above described *embodiments* of the invention.

12  
**CLAIMS**

1. A method of providing content to users via a communication infrastructure, wherein a plurality of content providers each make available content in XML format, and an portal  
5 operator provides a content-indexing portal through which users can access content from the content providers, the portal using XML style sheets to apply at least one of the transformations to XML-based content retrieved from a content provider in response to a user request:

- user device adaptation of the content;
- 10 - user preferences adaptation;
- location-based adaptation;

the content providers providing their content according to an agreed semantic meaning for the XML whereby to enable the portal to provide consistent user responses in respect of content from said plurality of content providers.

15

2. A method according to claim 1, wherein the semantic meaning of the XML is specified by the portal operator.

3. A method of providing content to users via a communication infrastructure, wherein a  
20 plurality of content providers each make available content in XML format, and an portal operator provides a content-aggregating portal through which users can access content from the content providers with the portal serving to aggregate content from multiple content providers as appropriate to provide a response to a user request, the portal using XML style sheets to apply at least one of the transformations to XML-based content  
25 retrieved from a content provider in response to a user request:

- user device adaptation of the content;
- user preferences adaptation;
- location-based adaptation;

the content providers providing their content according to an agreed semantic meaning for  
30 the XML whereby to enable the portal to provide consistent user responses when aggregating from multiple said content providers.

4. A method according to claim 3, wherein the agreed semantic meaning of the XML is specified by the portal operator.
- 5 5. A method according to claim 3, wherein the portal includes a two-level cache for caching both raw content obtained from the content providers and for caching content aggregated by the portal.
6. A method according to claim 3 or claim 5, wherein the portal aggregates and caches  
10 content in advance of any related request.
7. A method according to claim 6, wherein the content aggregation is effected in accordance with predetermined user content-personalisation profiles.
- 15 8. A method of providing services to users via a communication infrastructure, wherein a plurality of service providers each make available content in XML format associated with services, and an portal operator provides a portal through which users can access content from the service providers, the portal using XML style sheets to apply at least one of the transformations to XML-based content retrieved from a content provider in response to a  
20 user request:
- user device adaptation of the content;
  - user preferences adaptation;
  - location-based adaptation;
- the service providers providing their content according to an agreed semantic meaning for  
25 the XML whereby to enable the portal to provide consistent user responses in respect of services from said plurality of service providers.
9. A method according to claim 8, wherein the common XML semantic is specified in a registry run by the portal operator in association with actions associated with the related  
30 services.

10. A method according to claim 8, wherein the portal includes a two-level cache for caching both raw content obtained from the service providers and for caching content aggregated by the portal.
- 5 11. A method according to claim 8 or claim 10, wherein the portal aggregates and caches content in advance of any related request.
12. A method according to claim 8, wherein the content aggregation is effected in accordance with predetermined user service-personalisation profiles.



INVESTOR IN PEOPLE

**Application No:** GB 0115461.6  
**Claims searched:** 1-12

**Examiner:** Phil Osman  
**Date of search:** 14 March 2002

## Patents Act 1977 Search Report under Section 17

### Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:  
UK Cl (Ed.T):  
Int Cl (Ed.7): G06F 17/21, 17/22, 17/30  
Other: Online: EPODOC, JAPIO, WPI

### Documents considered to be relevant:

Category	Identity of document and relevant passage	Relevant to claims
Y	GB 2,346,238 A (IBM) See page 2 line 41 - page 3 line 16	1, 3, 8 at least
A, E	WO 01/77842 A1 (TELECOMMUNICATION SYSTEMS INC.) See Page 2 line 32 - page 3 line 30.	
A, E	WO 01/76190 A2 (WIRE-LESS KNOWLEDGE) See page 7 line 30 - page 8 line 14.	
A, E	WO 01/65355 A1 (CELL-TREX) See page 16 lines 2-6.	
Y	WO 01/42989 A2 (INTERLEAF) See page 1 line 30 - page 2 line 9 & page 4 lines 13-20.	1, 3, 8 at least
Y	WO 01/35235 A1 (VOCAL POINT) See page 12 lines 5-21	1, 3, 8 at least

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.